

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-209580

(43)Date of publication of application : 28.07.2000

(51)Int.Cl.

H04N 7/24

(21)Application number : 11-007035

(71)Applicant : CANON INC

(22)Date of filing : 13.01.1999

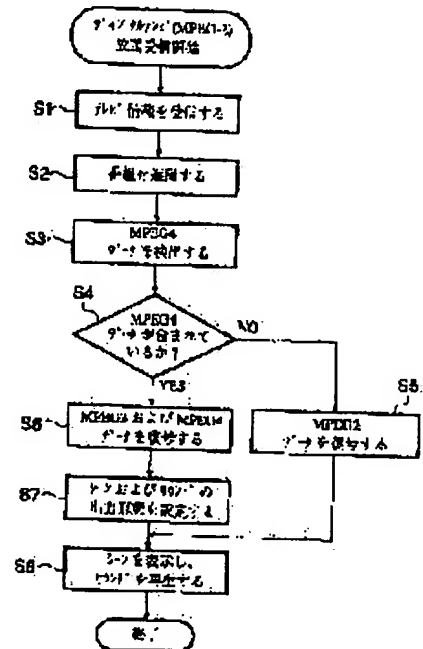
(72)Inventor : ITO MASAMICHI
TAKAHASHI KOJI

(54) PICTURE PROCESSOR AND ITS METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To acquire information desired by a user on each occasion as sub data with a small information amount ancillary to a main picture in a form of a picture independently of whether the information relates to or does not relate to the main picture in order to realize a television broadcast with many more functions.

SOLUTION: Digital television information in compliance with the MPEG2 is received (S1), a program is selected (S2), MPEG4 data of sub television information multiplexed on MPEG2 data stream of the selected program are detected (S3), and whether or not the MPEG4 data are included in the MPEG2 stream is discriminated based on the result of the detection (S4). When the MPEG4 data are included, the MPEG4 data are separated from the MPEG2 data stream, and the MPEG2 data and the MPEG4 data are respectively divided into picture data, sound data and system data and they are respectively decoded (S6). Moreover, an output form of the picture and the sound data in compliance with the MPEG2 and of a scene and the sound data in compliance with the MPEG4 can be set (S7).



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-209580
(P2000-209580A)

(43) 公開日 平成12年7月28日 (2000.7.28)

(51) Int.Cl.⁷
H 0 4 N 7/24

識別記号

F I
H 0 4 N 7/13

データベース (参考)
Z 5 C 0 5 9

審査請求 未請求 請求項の数9 O L (全 27 頁)

(21) 出願番号 特願平11-7035

(22) 出願日 平成11年1月13日 (1999.1.13)

(71) 出願人 000001007
キヤノン株式会社
東京都大田区下丸子3丁目30番2号
(72) 発明者 伊藤 賢道
東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内
(72) 発明者 ▲高▼橋 宏爾
東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内
(74) 代理人 100076428
弁理士 大塚 康徳 (外2名)

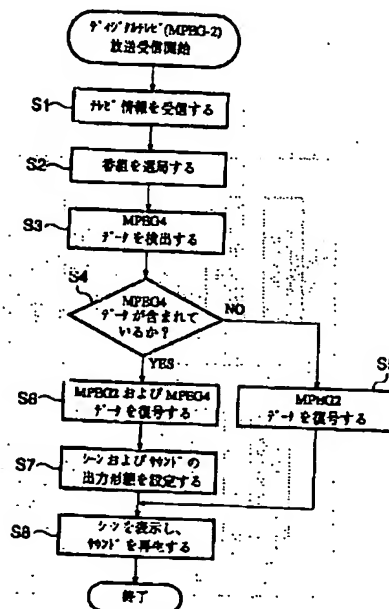
最終頁に続く

(54) 【発明の名称】 画像処理装置およびその方法

(57) 【要約】

【課題】 より多機能なテレビ放送などを実現する上で、メイン画像に関係する、または関係しなくとも、その時々によりユーザが欲する情報を画像の形態で、メイン画像に付随する小情報量のサブデータとして取得できることが望まれる。

【解決手段】 MPEG2のデジタルテレビ情報が受信され (S1)、番組が選局され (S2)、選局された番組のMPEG2データストリームに多重されているサブテレビ情報のMPEG4データの検出が行われ (S3)、その検出結果に基づき、MPEG2データストリームにMPEG4データが含まれているかが判断される (S4)。MPEG4データが含まれている場合、MPEG2データストリームからMPEG4データストリームが分離され、MPEG2データおよびMPEG4データそれぞれが画像、サウンドおよびシステムデータに分けられて復号される (S6)。さらに、MPEG2の画像およびサウンドと、MPEG4のシーンおよびサウンドとの出力形態が設定される (S7)。



【特許請求の範囲】

【請求項1】 MPEG2データストリームを入力する入力手段と、
 入力されるデータストリームからMPEG4データストリームを検出する検出手段と、
 入力されるデータストリームからMPEG2およびMPEG4で符号化された各データをそれぞれ分離する分離手段と、
 分離された各データを復号する復号手段と、
 前記検出手段による検出結果に基づき、復号された少なくとも画像情報の再生を制御する制御手段とを有することを特徴とする画像処理装置。

【請求項2】 前記入力手段には、デジタルテレビ放送のデータストリームが入力されることを特徴とする請求項1に記載された画像処理装置。

【請求項3】 前記入力手段には、データ蓄積メディアから再生されたデータストリームが入力されることを特徴とする請求項1に記載された画像処理装置。

【請求項4】 さらに、少なくとも画像情報の再生方法がマニュアル入力される指示入力手段を有し、
 入力された再生方法に基づき、前記制御手段による制御が行われることを特徴とする請求項1から請求項3の何れかに記載された画像処理装置。

【請求項5】 前記MPEG4データストリームには、さらにサウンド情報および/またはシステムデータが含まれ、前記制御手段は、復号されたシステムデータおよび入力された再生方法に基づき、画像および/またはサウンド情報の再生を制御することを特徴とする請求項4に記載された画像処理装置。

【請求項6】 前記MPEG4データストリームは、前記MPEG2データストリームのアダプテーションフィールドに多重されていることを特徴とする請求項1から請求項5の何れかに記載された画像処理装置。

【請求項7】 前記MPEG2データストリームに多重された前記MPEG4データストリームは、MPEG4未対応の装置には無視されることを特徴とする請求項1から請求項5の何れかに記載された画像処理装置。

【請求項8】 MPEG2データストリームを入力し、
 入力されたデータストリームからMPEG4データストリームを検出し、
 受信されたデータストリームからMPEG2およびMPEG4で符号化された各データをそれぞれ分離し、
 分離された各データを復号し、
 前記検出結果に基づき、復号された少なくとも画像情報の再生を制御することを特徴とする画像処理方法。

【請求項9】 画像処理のプログラムコードが記録された記録媒体であって、前記プログラムコードは少なくとも、
 MPEG2データストリームを入力するステップのコードと、
 入力されたデータストリームからMPEG4データストリー

ムを検出するステップのコードと、

受信されたデータストリームからMPEG2およびMPEG4で符号化された各データをそれぞれ分離するステップのコードと、

分離された各データを復号するステップのコードと、
 前記検出結果に基づき、復号された少なくとも画像情報の再生を制御するステップのコードとを有することを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は画像処理装置およびその方法に関し、例えば、MPEG2のデータストリームのようなデジタルデータ列から少なくとも画像を再生する画像処理装置およびその方法に関するものである。

【0002】

【従来の技術】近年、衛星放送やケーブル放送を用いたデジタルテレビ放送が開始された。デジタル放送の実現により、画像やサウンドを含むサウンドの品質向上、圧縮技術を利用した番組の種類や量の増大、インタラクティブサービスなど新しいサービスの提供、受信形態の進化など多くの効果が期待される。

【0003】図1は衛星放送を用いたデジタル放送受信機10の構成例を示すブロック図である。

【0004】まず、放送衛星によって送信されたテレビ(TV)情報がアンテナ1で受信される。受信されたTV情報はチューナ2で選局され復調される。その後、図示しないが、誤り訂正処理、必要であれば課金処理やデスクランブル処理などが行われる。次に、TV情報として多重化されている各種データを多重信号分離回路3で分離する。TV情報は画像情報、サウンド情報およびその他の付加データに分離される。分離された各データは復号回路4で復号される。こうして復号された各データのうち画像情報とサウンド情報はD/A変換回路5でアナログ化され、テレビジョン受信機(TV)6で再生される。一方、付加データは、番組サブデータとしての役割をもち各種機能に関与する。

【0005】さらに、受信されたTV情報の記録再生にはVTR7が利用される。受信機10とVTR7との間はIEEE1394などのデジタルインタフェースで接続されている。このVTR7は、デジタル記録方式の記録形態を備え、例えばD-VHS方式などによりTV情報をビットストリーム記録する。なお、D-VHS方式のビットストリーム記録に限らず、その他の民生用デジタル記録方式であるDVフォーマットや、DVD(Digital Video Disc)などの各種ディスク媒体を用いたデジタル記録装置などでもデジタルテレビ放送と同様のTV情報を記録することが可能である。ただし、フォーマット変換が必要になる場合がある。

【0006】また、上述したようなデジタルテレビ放送やデジタル記録装置では、主として、MPEG2で符号化されたデータフォーマットが採用されている。

【0007】

【発明が解決しようとする課題】しかしながら、地上波放送および上述したデジタルテレビ放送においてテレビ番組表をTV6に表示するには、放送局から送信されてくるメイン画像を単純に表示する方法だけしかない。メイン画像に付随するサブ情報を表示する例として文字放送があるが、提供される情報はテキスト情報など限られたものであり、画像を扱うことはできない。

【0008】また、複数チャネルの画像をマルチ画面で表示するテレビジョン受信機はあるが、それぞれの画像は、大情報量のメイン画像として送信されたものである。

【0009】より多機能なテレビ放送などを実現する上で、メイン画像に関係する、または関係しなくとも、その時々ユーザが欲する情報を画像（サウンドを加えてもよい）の形態で、メイン画像に付随する小情報量のサブデータとして取得できることが望まれる。しかし、このような技術は実現されていない。

【0010】本発明は、上述の問題を解決するためのものであり、メイン画像に関係する、または関係しなくとも、その時々所望される情報を少なくとも画像の形態で再生する機能を提供することを目的とする。

【0011】

【課題を解決するための手段】本発明は、前記の目的を達成する一手段として、以下の構成を備える。

【0012】本発明にかかる画像処理装置は、MPEG2データストリームを入力する入力手段と、入力されるデータストリームからMPEG4データストリームを検出する検出手段と、入力されるデータストリームからMPEG2およびMPEG4で符号化された各データをそれぞれ分離する分離手段と、分離された各データを復号する復号手段と、前記検出手段による検出結果に基づき、復号された少なくとも画像情報の再生を制御する制御手段とを有することを特徴とする。

【0013】本発明にかかる画像処理方法は、MPEG2データストリームを入力し、入力されたデータストリームからMPEG4データストリームを検出し、受信されたデータストリームからMPEG2およびMPEG4で符号化された各データをそれぞれ分離し、分離された各データを復号し、前記検出結果に基づき、復号された少なくとも画像情報の再生を制御することを特徴とする。

【0014】

【発明の実施の形態】〔概要〕本実施形態は、テレビ放送のメイン情報内の所定領域にサブ情報として、画像および/またはサウンドデータを含むサウンドデータを効率的に多重して送信し、受信側でメインおよびサブ情報を受信し再生にする。それぞれのデータフォーマットとしては、メイン画像情報にデジタルテレビ放送のMPEG2データストリームを、サブ情報に近年標準化が図られている、伝送効率が非常に高いMPEG4データストリーム

を用いる。

【0015】本実施形態によれば、メイン情報に多重されたサブ情報により画像やサウンドを伝送することができ、その時々ユーザが欲する情報を画像（サウンドを含むサウンドを加えてもよい）の形態で提供することが可能になる。さらに、視覚的効果を向上させることができる。

【0016】さらに、データフォーマットにMPEG2およびMPEG4を用いることで、現在のデジタルテレビ放送システムであるMPEG2との親和性も容易に実現できる。上、既存のMPEG2用コンテンツを有効に活用することができる。また、画像やサウンドをオブジェクト単位で扱うMPEG4は、サブ情報のデータフォーマットとして最適なデータ形態であると言える。

【0017】なお、本実施形態は、デジタルテレビ放送に限らず、DVD(Digital Video Disc)などのパッケージメディアにも適用することができる。

【0018】

【MPEG4の概要】〔規格の全体構成〕MPEG4規格は大きな四つの項目からなる。このうち三つの項目はMPEG2と類似していて、ビジュアルパート、オーディオパートおよびシステムパートである。

【0019】●ビジュアルパート

自然画、合成画、動画および静止画などを扱うオブジェクト符号化方式が規格として定められている。また、伝送路誤りの訂正や修復に適した符号化方式、同期再生機能および階層符号化が含まれている。表現上『ビデオ』は自然画像を意味し、『ビジュアル』は合成画像まで含む。

【0020】●オーディオパート

自然音、合成音および効果音などを対象にしたオブジェクト符号化方式が規格として定められている。ビデオパートやオーディオパートでは複数の符号化方式を規定し、それぞれのオブジェクトの特徴に適した圧縮方式を適宜選択することで、符号化効率を高める工夫がされている。

【0021】●システムパート

符号化された映像オブジェクトやサウンドオブジェクトの多重化処理と、その逆の分離処理を規定する。さらにバッファメモリや時間軸制御と再調整機能もこのパートに含まれる。上記のビジュアルパートおよびオーディオパートで符号化された映像オブジェクトやサウンドオブジェクトは、シーンのなかのオブジェクトの位置や出現時刻および消滅時刻などを記したシーン構成情報とともにシステムパートの多重化ストリームに統合される。復号処理としては、受信したビットストリームから、それぞれのオブジェクトを分離/復号し、シーン構成情報に基づきシーンを再構成する。

【0022】〔オブジェクトの符号化〕MPEG2ではフレームあるいはフィールドを単位として符号化するが、コ

コンテンツの再利用や編集を実現するために、MPEG4では映像データやオーディオデータをオブジェクト（物体）として扱う。オブジェクトには以下のような種類がある。

サウンド

自然画像（背景映像：二次元固定映像）

自然画像（主被写体映像：背景なし）

合成画像

文字画像

【0023】これらを同時に入力し符号化処理する場合のシステム構成例を図2に示す。サウンドオブジェクト符号化器5001、自然画像オブジェクト符号化器5002、合成画像オブジェクト符号化器5003および文字オブジェクト符号化器5004により、オブジェクトはそれぞれ符号化処理される。この符号化と略同時に、シーン内における各オブジェクトの関連を示すシーン構成情報を、シーン記述情報符号化器5005により符号化する。符号化されたオブジェクト情報およびシーン記述情報は、データ多重化器5006によりMPEG4ビットストリームへエンコード処理される。

【0024】このようにエンコード側では、複数のビジュアルオブジェクトやオーディオオブジェクトの組み合わせを定義して、一つのシーン（画面）を表現する。ビジュアルオブジェクトに関しては、自然画像とコンピュータグラフィクスなどの合成画像とを組み合わせたシーンも構成できる。また、上記のような構成をとることで、例えば、テキストサウンド合成の機能を使って、被写体映像とそのサウンドとの同期再生が可能になる。なお、前記のビットストリーム状態で送受信または記録再生が行われる。

【0025】デコード処理は、先のエンコード処理の逆処理である。データ分離器5007により、MPEG4ビットストリームを各オブジェクトに分離し分配する。分離されたサウンド、自然画像、合成画像および文字などのオブジェクトは、対応する復号器5008から5011によりオブジェクトデータへ復号される。また、シーン記述情報も略同時に復号器5012により復号される。これらの復号情報を用いて、シーン合成器5013は、元のシーンを合成する。

【0026】デコード側では、シーンに含まれるビジュアルオブジェクトの位置や、オーディオオブジェクトの順番など、部分的な変更が可能である。オブジェクト位置はドラッグにより変更でき、言語の変更などはユーザがオーディオオブジェクトを変更することで可能になる。

【0027】複数のオブジェクトを自由に組み合わせてシーンを合成するために、次の四つの項目が規定されている。

【0028】●オブジェクト符号化

ビジュアルオブジェクトおよびオーディオオブジェク

ト、並びに、それらを組み合わせたAV（オーディオビジュアル）オブジェクトを符号化対象とする。

【0029】●シーン合成

ビジュアルオブジェクト、オーディオオブジェクトおよびAVオブジェクトを所望するシーンに構成するためのシーン構成情報と合成方式とを規定するために、Virtual Reality Modeling Language (VRML) をモディファイした言語を用いる。

【0030】●多重化と同期

各オブジェクトを多重同期したストリーム（エレメンタリストリーム）の形式などを定める。このストリームをネットワークに流したり、記録装置に格納するときのサービス品質QOS (Quality of Service) も設定可能である。QOSパラメータには、最大伝送速度、誤り率および伝送方式などの伝送路条件や復号能力などが設けられている。

【0031】●ユーザの操作（インタラクション）

ビジュアルオブジェクトやオーディオオブジェクトを、ユーザ端末側で合成する方式を定義する。MPEG4のユーザ端末は、ネットワークや記録装置から送られてくるデータを、エレメンタリストリームに分離して、各オブジェクトごとに復号する。複数の符号化されたデータから、同時に送られてきたシーン構成情報を基にしてシーンを再構成する。

【0032】ユーザ操作（編集）を考慮に入れたシステムの構成例を図3に示す。また、ビデオオブジェクトに関するVOP処理回路のエンコーダ側のブロック図を図4に、デコーダ側のブロック図を図5に示す。

【0033】[VOP (Video Object Plane)] MPEG4における映像の符号化は、対象の映像オブジェクトを形状 (Shape) とその絵柄 (Texture) に分けてそれぞれ符号化する。この映像データの単位をVOPという。図6はVOPの符号化および復号の全体構成を示すブロック図である。

【0034】例えば、画像が人物と背景の二つのオブジェクトから構成されている場合、各フレームを二つのVOPに分割して符号化する。各VOPを構成する情報は、図7Aに示されるように、オブジェクトの形状情報、動き情報およびテキスト情報である。一方、復号器は、ビットストリームをVOP毎に分離し個別に復号した後、これらを合成して画像を形成する。

【0035】このように、VOP構造の導入により、処理対象の画像が複数の映像オブジェクトから構成されている場合、これを複数のVOPに分割し、個別に符号化/復号することができる。なお、VOPの数が1で、オブジェクト形状が矩形の場合は、図7Bに示すように、従来からのフレーム単位の符号化になる。

【0036】VOPには三種類の予測方式、面内符号化 (I-VOP)、前方向予測 (P-VOP) および双方向予測 (B-VOP) がある。予測単位は16×16画素のマクロブロック (MB) である。

【0037】双方向予測B-VOPは、MPEG1およびMPEG2のBピクチャと同じく、過去のVOPおよび未来のVOPの両方向からVOPを予測する方式である。そして、マクロブロック単位に直接符号化/前方符号化/後方符号化/双方符号化の四種類のモードが選択可能である。そしてこのモードは、MBまたはブロック単位に切り替えることが可能である。P-VOPの動きベクトルのスケーリングで双方向予測する。

【0038】[形状(Shape)符号化]オブジェクト(物体)単位で画像を扱うためには、物体の形状が符号化および復号の際に既知でなければならない。また、後方にある物体が透けて見えるガラスのような物体を表現するためには、物体の透明度を表す情報が必要になる。この物体の形状および物体の透明度の情報を合わせて形状情報と呼ぶ。そして、形状情報の符号化を形状符号化と呼ぶ。

【0039】[サイズ変換処理]二値形状符号化は、画素毎に物体の外側にあるのか内側にあるのかを判定して、境界画素を符号化する手法である。従って、符号化すべき画素数が少ないほど発生符号量も少なくて済む。しかし、符号化すべきマクロブロックサイズを小さくすることは、元の形状符号が劣化して受信側に伝送されることになる。従って、サイズ変換により元の情報がどの程度劣化するかを測定し、所定のしきい値以下のサイズ変換誤差が得られる限りは、できるだけ小さなマクロブロックサイズを選択する。具体的なサイズ変換比率としては、原寸大、縦横1/2倍、縦横1/4倍の三種類が挙げられる。

【0040】各VOPの形状情報は、8ビットの α 値として与えられ、次のように定義される。

$\alpha = 0$: 該当VOPの外側

$\alpha = 1 \sim 254$: 他のVOPと半透明状態で表示

$\alpha = 255$: 該当VOPのみの表示領域

【0041】二値形状符号化は、 α 値が0あるいは255をとる場合であり、該当VOPの内側と外側のみで形状が表現される。多値形状符号化は、 α 値が0から255のすべての値を取り得る場合で、複数のVOP同士が半透明で重畳された状態を表現することができる。

【0042】テキストチャ符号化と同様に16×16画素のブロック単位に一画素精度の動き補償予測をする。オブジェクト全体を面内符号化する場合は形状情報の予測はしない。動きベクトルは、隣接するブロックから予測した動きベクトルの差分を用いる。求めた動きベクトルの差分値は、符号化してからビットストリームに多重化する。MPEG4では、動き補償予測したブロック単位の形状情報を二値形状符号化する。

【0043】●フェザリング

その他、二値形状の場合でも、境界部を不透明から透明に滑らかに変化させたい場合はフェザリング(境界形状のスムージング)を使う。フェザリングは、境界値

を線形に補間する線形フェザリングモードと、フィルタを使うフェザリングフィルタモードがある。不透明度が一定な多値形状には、定アルファモードがあり、フェザリングと組み合わせが可能である。

【0044】[テキストチャ符号化]オブジェクトの輝度成分や色差成分の符号化を行うもので、フィールド/フレーム単位のDCT(Discrete Cosine Transfer)、量子化、予測符号化および可変長符号化の順に処理する。

【0045】DCTは8×8画素のブロックを処理単位とするが、オブジェクト境界がブロック内にある場合は、オブジェクトの平均値でオブジェクト外の画素を補填する。その後、4タップの二次元フィルタ処理を施すことで、DCT係数に大きな擬似ピークが発生する現象を防ぐ。

【0046】量子化はITU-T勧告H.263の量子化器あるいはMPEG2の量子化器の何れかを使う。MPEG2量子化器を使えば、直流成分の非線形量子化やAC成分の周波数重み付けが可能になる。

【0047】量子化後の面内符号化係数は、可変長符号化する前にブロック間で予測符号化し冗長成分を削除する。とくに、MPEG4では直流成分と交流成分の両方に対して予測符号化する。

【0048】テキストチャ符号化のAC/DC予測符号化は、図8に示すように、注目ブロックに隣接するブロック間で対応する量子化係数の差分(勾配)を調べ、小さい方の量子化係数を予測に使う。例えば、注目ブロックの直流係数 x を予測する場合、対応する隣接ブロックの直流係数が a 、 b および c ならば次のようになる。

$|a - b| < |b - c|$ ならば直流係数 c を予測に使う

$|a - b| \geq |b - c|$ ならば直流係数 a を予測に使う

【0049】注目ブロックの交流係数 X を予測する場合も、上記と同様に予測に使う係数を選んだ後、各ブロックの量子化スケール値 QP で正規化する。

【0050】直流成分の予測符号化は、隣接するブロック間で上下に隣接するブロックの直流成分の差(垂直勾配)と、左右に隣接するブロックの直流成分の差(水平勾配)を調べ、勾配の少ない方向のブロックの直流成分との差分を予測誤差として符号化する。

【0051】交流成分の予測符号化は、直流成分の予測符号化に合わせて、隣接ブロックの対応する係数を用いる。ただし、量子化パラメータの値がブロック間で異なっている可能性があるため、正規化(量子化ステップスケーリング)してから差分をとる。予測の有無はマクロブロック単位に選択できる。

【0052】その後、交流成分は、ジグザグスキャンされ、三次元(Last、RunおよびLevel)可変長符号化される。ここで、Lastはゼロ以外の係数の終りを示す1ビットの値、Runはゼロの継続長、Levelは非ゼロ係数の値である。

【0053】面内符号化された直流成分の可変長符号化

には、直流成分用可変長符号化テーブルまたは交流成分用可変長テーブルの何れかを使う。

【0054】〔動き補償〕MPEG4では任意の形状のビデオオブジェクトブレン(VOP)を符号化することができる。VOPには、前述したように、予測の種類によって面内符号化(I-VOP)、前方向予測符号化(P-VOP)および双方向予測符号化(B-VOP)があり、予測単位は16ライン×16画素または8ライン×8画素のマクロブロックを使う。従って、VOPの境界上に跨るマクロブロックも存在することになる。このVOP境界の予測効率を改善するために、境界上のマクロブロックに対してはパディング(補填)およびポリゴンマッチング(オブジェクト部分のみのマッチング)を行う。

【0055】〔ウェーブレット符号化〕ウェーブレット(wavelet)変換は、一つの孤立波関数を拡大/縮小/平行移動して得られる複数の関数を変換基底とする変換方式である。このウェーブレット変換を用いた静止画像の符号化モード(Texture Coding Mode)は、とくにコンピュータグラフィックス(CG)画像と自然画像とが合成された画像を扱う場合に、高解像度から低解像度までの様々な空間解像度を備えた高画質の符号化方式として適している。ウェーブレット符号化は、画像をブロック分割せず一括して符号化することができるため、低ビットレートでもブロック歪みが発生せず、モスキート雑音も減少できる。このように、MPEG4の静止画像符号化モードは、低解像度かつ低画質の画像から高解像度かつ高画質の画像までの幅広いスケーラビリティ、処理の複雑性および符号化効率のトレードオフの関係をアプリケーションに応じて調整できる。

【0056】〔階層符号化(スケーラビリティ)〕スケーラビリティを実現するために、図9Aおよび9Bに示すようなシンタックスの階層構造を構成する。階層符号化は、例えばベースレイヤを下位レイヤ、補強レイヤを上位レイヤとし、補強レイヤにおいてベースレイヤの画質を向上する「差分情報」を符号化することによって実現される。空間スケーラビリティの場合、ベースレイヤは低解像度の動画像を、「ベースレイヤ+補強レイヤ」で高解像度の動画像を表す。

【0057】さらに、画像全体の画質を階層的に向上させるほかに、画像中の物体領域のみ画質を向上させる機能がある。例えば、時間スケーラビリティの場合、ベースレイヤは画像全体を低いフレームレートで符号化したもの、補強レイヤは画像内の特定オブジェクトのフレームレートを向上させるデータを符号化したものになる。

【0058】●時間スケーラビリティ

図9Aに示す時間スケーラビリティは、フレーム速度を階層化し、補強レイヤのオブジェクトのフレーム速度を速くすることができる。階層化の有無はオブジェクト単位で設定できる。補強レイヤのタイプは二つで、タイプ1はベースレイヤのオブジェクトの一部で構成する。タイ

プ2はベースレイヤと同じオブジェクトで構成する。

【0059】●空間スケーラビリティ

図9Bに示す空間スケーラビリティは空間解像度を階層化する。ベースレイヤは、任意のサイズのダウンサンプリングが可能で、補強レイヤの予測に使用される。

【0060】〔スプライト符号化〕スプライトとは、三次元空間画像における背景画像などのように、オブジェクト全体が統一的に移動、回転、変形などを表現できる平面的なオブジェクトのことである。この平面的オブジェクトを符号化する手法をスプライト符号化と呼ぶ。

【0061】スプライト符号化は四種、静的/動的およびオンライン/オフラインに区別される。詳しく説明すると、オブジェクトデータを予め復号器に送り、グローバル動き係数だけをリアルタイムに伝送する構成であって、テンプレートオブジェクトの直接変換で得られる静的スプライト。時間的に前のスプライトからの予測符号化により得られる動的スプライト。事前に面内符号化(I-VOP)により符号化され、復号器側に伝送されるオフラインスプライト。符号化中に符号化器および復号器で同時に作成されるオンラインスプライトがある。

【0062】スプライト符号化に関して検討されている技術には、スタティックスプライト(Static Sprite)符号化、ダイナミックスプライト(Dynamic Sprite)符号化、グローバル動き補償などがある。

【0063】●スタティックスプライト符号化

スタティックスプライト符号化は、ビデオクリップ全体の背景(スプライト)を予め符号化しておき、背景の一部を幾何変換することによって画像を表現する方法である。切り出された一部の画像は、平行移動、拡大/縮小、回転など様々な変形を表現することができる。これについて図10Aに示すように、画像の移動、回転、拡大/縮小、変形などにより三次元空間における視点移動を表現することをワーブと呼ぶ。

【0064】ワーブの種類には遠近法変換、アフィン変換、等方拡大(a)/回転(θ)/移動(c, f)および平行移動があり、図10Bの各式で表される。図10Bに示す式の係数によって移動、回転、拡大/縮小、変形などが表される。また、スプライトの生成は符号化開始前にオフラインで行われる。

【0065】このように、背景画像の一部領域を切り取り、この領域をワーブして表現することでスタティックスプライト符号化は実現される。図11に示すスプライト(背景)画像に含まれる一部領域がワーブされることになる。例えば、背景画像はテニスの試合における観客席などの画像であり、ワーブされる領域はテニスプレーヤなどの動きのあるオブジェクトを含んだ画像である。また、スタティックスプライト符号化においては、幾何変換パラメータのみを符号化して、予測誤差を符号化しない。

【0066】●ダイナミックスプライト符号化

スタティックスプライト符号化では符号化前にスプライトが生成される。これに対して、ダイナミックスプライト符号化では、符号化しながらオンラインにスプライトを更新することができる。また、予測誤差を符号化するという点でスタティックスプライト符号化とは異なる。

【0067】●グローバル動き補償(GMC)

グローバル動き補償とは、オブジェクト全体の動きを、ブロックに分割することなく、一つの動きベクトルで表して動き補償する技術であり、剛体の動き補償などに適している。参照画像が、スプライトの代わりに直前の復号画像になる点、予測誤差を符号化する点では、スタティックスプライト符号化と同様である。ただし、スプライトを格納するためのメモリを必要としないこと、形状情報が不要であることは、スタティックスプライト符号化およびダイナミックスプライト符号化とは異なる。画面全体の動きや、ズームを含む画像などにおいて効果がある。

【0068】[シーン構造記述情報] シーン構成情報によりオブジェクトは合成される。MPEG4では、各オブジェクトをシーンに合成するための構成情報を伝送する。個別に符号化された各オブジェクトを受信したときに、シーン構成情報を使えば、送信側が意図したとおりのシーンに合成できる。

【0069】シーン構成情報には、オブジェクトの表示時間や表示位置などが含まれ、これらがツリー状のノード情報として記述されている。各ノードは、親ノードに対する時間軸上の相対時刻情報と相対空間座標位置情報をもつ。シーン構成情報を記述する言語には、VRMLを修正したBIFS(Binary Format for Scenes)とJava(TM)を用いたAAVS(Adaptive Audio-Visual Session Format)がある。BIFSは、MPEG4のシーン構成情報を二値で記述する形式。AAVSはJava(TM)をベースとし、自由度が大きくBIFSを補う位置付けにある。図12はシーン記述情報の構成例を示す図である。

【0070】[シーン記述] シーン記述はBIFSによって行われる。ここでは、VRMLとBIFS共通の概念であるシーングラフとノードを中心に説明する。

【0071】ノードは光源、形状、材質、色および座標などの属性や、座標変換を伴う下位ノードのグループ化を指定する。オブジェクト指向の考えを取り入れ、三次元空間中の各物体の配置や見え方は、シーングラフと呼ばれる木を、頂点のノードから辿り、上位ノードの属性を継承することにより決定される。葉にあたるノードにメディアオブジェクト、例えば、MPEG4ビデオのビットストリームを同期をとって割当てれば、他のグラフィクスと共に動画を三次元空間内に合成して表示することができる。

【0072】また、VRMLとの差異は下記のとおりである。MPEG4システムでは次をBIFSでサポートする。

(1) MPEG4ビデオVOP符号化の二次元オーバラップ関係記

述と、MPEG4オーディオの合成記述

(2) 連続メディアストリームの同期処理

(3) オブジェクトの動的振る舞い表現(例えばスプライト)

(4) 伝送形式(バイナリ)を標準化

(5) セッション中にシーン記述を動的に変更

【0073】VRMLのノードのうちExtrusion、Script、ProtoおよびExtmProtoなどがサポートされていない以外は、VRMLノードのほぼすべてがBIFSでサポートされている。BIFSで新たに加えられたMPEG4特別ノードには、以下のものがある。

(1) 2D/3D合成のためのノード

(2) 2Dグラフィクスやテキストのためのノード

(3) アニメーションノード

(4) オーディオノード

【0074】特筆すべきは、VRMLでは背景など特殊なノードを除き2D合成はサポートされていなかったが、BIFSでは、テキストやグラフィックオーバーレイ、さらにMPEG4ビデオVOP符号化を画素単位で扱えるように記述が拡張されている。

【0075】アニメーションノードには、3Dメッシュで構成された顔などMPEG4のCG画像のための特別なノードが規定されている。シーングラフ中のノードの置き換え、消去、追加および属性変更が動的に行えるメッセージ(BIFS Update)があり、セッションの途中で画面上に新たな動画を表示したり、ボタンを追加することが可能になる。BIFSは、VRMLの予約語、ノード識別子および属性値をほぼ一対一にバイナリデータに置き換えることにより実現できる。

【0076】[MPEG4オーディオ] 図13にMPEG4オーディオの符号化方式の種類を示す。オーディオおよびサウンドの符号化には、パラメトリック符号化、CELP(Code Excited Linear Prediction)符号化、時間/周波数変換符号化が含まれる。さらに、SNHC(Synthetic Natural Hybrid Coding)オーディオの機能も取り入れ、SA(Structured Audio: 構造化オーディオ)符号化とTTS(Text to Speech: テキストサウンド合成)符号化が含まれる。SAはMIDI(Music Instrument Digital Interface)を含む合成楽音の構造的記述言語であり、TTSは外部のテキストサウンド合成装置にイントネーションや音韻情報などを送るプロトコルである。

【0077】図14にオーディオ符号化方式の構成を示す。図14において、入力サウンド信号を前処理(201)し、パラメトリック符号化(204)、CELP符号化(205)および時間/周波数符号化(206)の三つの符号化を使い分けるように、帯域に応じて信号分割(202)し、それぞれに適した符号化器へ入力する。また、信号分析制御(203)により、入力オーディオ信号が分析され、入力オーディオ信号を各符号化器へ割り当てるための制御情報などが発生される。

【0078】続いて、それぞれ別の符号化器であるパラメトリック符号化コア(204)、CELP符号化コア(205)、時間/周波数変換符号化コア(206)は、各符号化方式に基づいた符号化処理を実行する。これら三種の符号化方式については後述する。パラメトリック符号化およびCELP符号化されたオーディオデータは、小ステップ強化(207)され、時間/周波数変換符号化および小ステップ強化されたオーディオデータは、大ステップ強化(208)される。なお、小ステップ強化(207)および大ステップ強化(208)は、各符号化処理で発生する歪を減少させるためのツールである。こうして、大ステップ強化されたオーディオデータは、符号化されたサウンドビットストリームになる。

【0079】以上が図75のサウンド符号化方式の構成の説明であるが、次に、図13を参照しながら各符号化方式について説明する。

【0080】●パラメトリック符号化

サウンド信号や楽音信号を含むサウンド信号を周波数、振幅およびピッチなどのパラメータで表現し、それを符号化する。サウンド信号用の調波ベクトル駆動符号化(HVXC: Harmonic Vector Excitation Coding)と、楽音信号用の個別スペクトル(IL: Individual Line)符号化が含まれる。

【0081】HVXC符号化は、主として2k~4kbpsのサウンド符号化を目的とし、サウンド信号を有声音と無声音に分類し、有声音は線形予測係数(LPC: Linear Prediction Coefficient)の残差信号の調波(ハーモニック)構造をベクトル量子化する。無声音については、予測残差をそのままベクトル駆動符号化(vector excitation coding)する。

【0082】IL符号化は、6k~16kbpsの楽音の符号化を目的としており、信号を線スペクトルでモデル化して符号化するものである。

【0083】●CELP符号化

入力サウンド信号をスペクトル包絡情報と音源情報(予測誤差)とに分離して符号化する方式である。スペクトル包絡情報は、入力サウンド信号から線形予測分析によって算出される線形予測係数によって表される。MPEG4のCELP符号化には帯域幅4kHzの狭帯域CELPと、帯域幅8kHzの広帯域CELPがあり、狭帯域(NB: Narrow Band) CELPは3.85~12.2kbps、広帯域(WB: Wide Band) CELPは13.7k~24kbpsの間においてビットレートの選択が可能である。

【0084】●時間/周波数変換符号化

高音質を目指す符号化方式である。AAC(Advanced Audio Coding)に準拠する方式、およびTwinVQ(Transform-domain Weighted Interleave Vector Quantization: 変換領域重み付けインタリーブベクトル量子化)がこれに含まれる。この時間/周波数変換符号化には聴覚心理モデルが組み込まれ、聴覚マスキング効果を利用しながら適

応量子化する仕組みになっている。

【0085】AAC準拠方式は、オーディオ信号をDCTなどで周波数変換し、聴覚マスキング効果を利用しながら適応量子化する仕組みである。適応ビットレートは24k~64kbpsである。

【0086】TwinVQ方式は、オーディオ信号を線形予測分析したスペクトル包絡を用いて、オーディオ信号のMDCT係数を平坦化する。インタリーブを施した後、二つの符号長を用いてベクトル量子化する仕組みである。適応ビットレートは6k~40kbpsである。

【0087】[システム構造] MPEG4のシステムパートでは、多重化、分離および合成(コンポジション)を定義する。以下、図15を用いてシステム構造を説明する。

【0088】多重化においては、映像符号化器やオーディオ符号化器からの出力である各オブジェクトや、各オブジェクトの時空間配置を記述したシーン構成情報などのエレメンタリストリームごとに、アクセスユニットレイヤでパケット化される。アクセスユニットレイヤでは、アクセスユニット単位に同期を取るためのタイムスタンプや参照クロックなどがヘッダとして付加される。パケット化されたストリームは、次に、FlexMuxレイヤで表示や誤り耐性の単位で多重化され、TransMuxレイヤへ送られる。

【0089】TransMuxレイヤでは、誤り耐性の必要度に応じて誤り訂正符号が保護サブレイヤで付加される。最後に、多重サブレイヤ(Mux Sub Layer)で一本のTransMuxストリームとして伝送路に送り出される。TransMuxレイヤは、MPEG4では定義されず、インターネットのプロトコルであるUDP/IP(User Datagram Protocol/Internet Protocol)やMPEG2のトランスポートストリーム(TS)、ATM(Asynchronous Transfer Mode)のAAL2(ATM Adaptation layer2)、電話回線利用のテレビ電話用多重化方式(ITU-T勧告H.223)、および、デジタルオーディオ放送などの既存のネットワークプロトコルが利用可能である。

【0090】システムレイヤのオーバーヘッドを軽くし、従来のトランスポートストリームに容易に埋め込めるように、アクセスユニットレイヤやFlexMuxレイヤをバイパスすることも可能である。

【0091】復号側では、各オブジェクトの同期を取るために、デマルチプレクス(分離)の後段にバッファ(DB: Decoding Buffer)を設け、各オブジェクトの到達時刻や復号時間のずれを吸収する。合成の前にもバッファ(CB: Composition Buffer)を設けて表示タイミングを調整する。

【0092】[ビデオストリームの基本構造] 図16にレイヤ構造を示す。各階層をクラスと呼び、各クラスにはヘッダが付く。ヘッダとはstart code、end code、ID、形状およびサイズほかの各種符号情報である。

【0093】●ビデオストリーム

ビデオストリームは複数のセッションで構成される。セッションとは、一連の完結したシーケンスのことである。

VS: セッションは複数のオブジェクトで構成される

VO: ビデオオブジェクト

VOL: オブジェクトは複数のレイヤを含むオブジェクト単位のシーケンス

GOV: オブジェクトは複数のレイヤで構成される

VOP: オブジェクトレイヤは複数のプレーンで構成される

ただし、プレーンはフレーム毎のオブジェクト

【0094】[誤り耐性を有するビットストリーム構造] MPEG4は、移動体通信(無線通信)などに対応すべく、符号化方式自体が伝送誤りに対する耐性を有している。既存の標準方式における誤り訂正は主にシステム側で行っているが、PHS(Personal Handy phone System)などのネットワークでは誤り率が非常に高く、システム側では訂正しきれない誤りがビデオ符号化部分に漏れ込んでくることが予想される。これを考慮して、MPEG4は、システム側で訂正しきれなかった各種のエラーパターンを想定し、このような環境の下でも可能な限り誤りの伝播が抑制されるような誤り耐性符号化方式とされている。ここでは、画像符号化に関する誤り耐性の具体的な手法と、そのためのビットストリーム構造を説明する。

【0095】●Reversible VLC(RVLC)と双方向復号

図17に示すように、復号途中で誤りの混入が確認された場合、そこで復号処理を一旦停止し、次の同期信号の検出を行う。次の同期信号が検出できた段階で、今度はそこから逆向きにビットストリームの復号処理を行う。新たな付加情報なしに、復号のスタートポイントが増加していることになり、誤り発生時に復号できる情報量を従来よりも増やすことが可能になる。このような順方向と同時に逆方向からも復号可能な可変長符号により「双方向復号」が実現される。

【0096】●重要情報の複数回伝送

図18に示すように、重要情報を複数回伝送することが可能な構成を導入し、誤り耐性を強化する。例えば、各VOPを正しいタイミングで表示するためにはタイムスタンプが必要であり、この情報は最初のビデオパッケージに含まれている。仮に、誤りによってこのビデオパッケージが消失しても、前記の双方向復号構造により次のビデオパッケージから復号が再開できるが、このビデオパッケージにはタイムスタンプがないため、結局、表示タイミングがわからないことになる。そのため各ビデオパッケージにHEC(Header Extension Code)というフラグを立て、この後にタイムスタンプなどの重要情報を付加できる構造が導入された。HECフラグの後には、タイムスタンプとVOPの符号化モードタイプとが付加できる。

【0097】同期はずれが生じた場合は、次の同期回復マーカ(RM)から復号が開始されるが、各ビデオパッケージ

にはそのために必要な情報、そのパッケージに含まれる最初のMBの番号およびそのMBに対する量子化ステップサイズがRM直後に配置されている。その後にHECフラグが挿入され、HEC= '1' の場合にはTRおよびVCTがその直後に付加される。これらHEC情報により、仮に、先頭のビデオパッケージが復号できずに廃棄されても、HEC= '1' と設定したビデオパッケージ以降の復号および表示は正しく行われることになる。なお、HECを '1' にするか否かは符号化側で自由に設定できる。

【0098】●データパーティショニング

符号化側では、MB単位の符号化処理を繰り返してビットストリームを構成するため、途中に誤りが混入すると、それ以降のMBデータは復号できない。一方、複数のMB情報をまとめて幾つかのグループに分類し、それぞれをビットストリーム内に配置し、各グループの境目にマーカ情報を組み込めば、仮にビットストリームに誤りが混入してそれ以降のデータが復号できない場合でも、そのグループの最後にあるマーカで同期を取り直して、次のグループのデータを正しく復号することが可能になる。

【0099】以上の考えに基づき、ビデオパッケージ単位に、動きベクトルとテクスチャ情報(DCT係数など)とにグループ分けするデータパーティショニング手法(Data Partitioning)が採用されている。また、グループの境目にはモーションマーカ(MM:Motion Marker)が配置される。

【0100】仮に、動きベクトル情報の途中に誤りが混入していても、MMの後にくるDCT係数は正しく復号できるため、誤り混入以前の動きベクトルに対応するMBデータはDCT係数とともに正確に再生できる。またTexture部分に誤りが混入した場合でも、動きベクトルが正しく復号されていれば、その動きベクトル情報と復号済みの前フレーム情報とを用いて、ある程度正確な画像が補間再生(コンシールメント)できる。

【0101】●可変長間隔同期方式

ここでは、可変長パッケージで構成されている同期回復手法を説明する。先頭に同期信号を含んだMB群は「ビデオパッケージ」と呼ばれ、その中に何個のMBを含めるかは符号化側で自由に設定することができる。可変長符号(VLC: Variable Length Code)を使用するビットストリームに誤りが混入した場合、それ以降の符号の同期が取れなくなり、復号不可能な状態になる。このような場合でも、次の同期回復マーカを検出することにより、その後の情報を正しく復号することが可能になる。

【0102】[バイトアライメント] システムとの整合性をとるために、情報の多重化はバイトの整数倍単位で行われる。ビットストリームは、バイトアラインメント(Byte alignment)構造となっている。バイトアラインメントを行うために、各ビデオパッケージの最後にスタッフビットが挿入される。さらにこのスタッフビットは、ビデオパッケージ内のエラーチェック符号としても使用され

る。

【0103】スタッフビットは '0111' のように、最初の1ビットが '0' で、それ以外のビットがすべて '1' であるような符号で構成されている。つまりビデオパケット内の最後のMBまで正しく復号されれば、その次に来る符号は必ず '0' であり、その後はスタッフビットの長さより1ビット分だけ短い '1' の連続があるはずである。もし、このルールに反したパターンが検出された場合、それ以前の復号が正しく行われていないことになり、ビットストリームに誤りが混入していたことが検出できる。

【0104】以上、「国際標準規格MPEG4の概要決まる」(日経エレクトロニクス 1997.9.22号 p.147-168)、「見えてきたMPEG4の全貌」(社団法人映像情報メディア学会テキスト 1997.10.2)、「MPEG4の最新標準化動向と画像圧縮技術」(日本工業技術センター セミナー資料 1997.2.3)などを参考にして、MPEG4の技術に関して説明した。

【0105】

【第1実施形態】〔構成〕以下、本発明にかかる一実施形態のテレビ放送受信装置を図面を参照して詳細に説明する。図19は本発明にかかる実施形態のテレビ放送受信装置の構成例を示すブロック図である。

【0106】デジタルテレビ放送の信号は、その放送形態に応じて、衛星放送のときは衛星アンテナ21およびチューナ23により、ケーブル放送のときはケーブル22を介してチューナ24により、選局され受信される。こうして衛星放送もしくはケーブル放送から受信されたテレビ情報は、データ選択器43により一方のデータ列が選択され、復調回路25で復調され、誤り訂正回路26で誤り訂正される。

【0107】IEEE1394などのデジタルインタフェースをサポートするデータインタフェース(DIF)54を介して接続されたDVDやVTRなどの記録再生装置へ誤り訂正されたテレビ情報を記録したり、記録再生装置により再生されたテレビ情報を入力することができる。

【0108】次に、MPEG4データ検出回路51は、誤り訂正されたテレビ情報のデータ列内にMPEG4データが含まれているかを検出するものである。

【0109】本実施形態におけるテレビ情報とは、MPEG2で符号化されたテレビ情報であるメインの画像およびサウンドに、MPEG4で符号化された小データ量の画像および/またはサウンドのオブジェクトを多重した形態を有する。従って、MPEG4データ検出回路51は、主としてテレビ情報を構成するMPEG2データストリーム内の所定領域に、サブデータとしてのMPEG4データストリームが含まれているかを検出するものである。勿論、MPEG2データストリームに付加されたMPEG4データストリームの存在を示す識別用のIDなどを検出する方法も、検出方法の一つである。なお、MPEG2データストリームにMPEG4デ

ータストリームを多重する方法の詳細は後述する。

【0110】MPEG4データ検出回路29は、MPEG2データストリーム内にMPEG4オブジェクトが多重されていることを検出した場合、システムコントローラ38へその旨を示す信号を送る。システムコントローラ38は、その信号に応じて、画像やサウンドの再生表示を制御する。

【0111】一方、テレビ情報は、多重データ分離回路27により、MPEG2サウンド、MPEG2画像およびMPEG2システムデータがそれぞれの復号回路に合わせて分離される。さらに、MPEG2データストリームにMPEG4データストリームが含まれる場合、多重データ分離回路27により、MPEG4サウンドオブジェクト、MPEG4画像オブジェクトおよびシーン記述情報などを含むMPEG4システムデータがそれぞれの復号回路に合わせて分離される。

【0112】分離された各データまたはオブジェクトは、MPEG2サウンド復号回路28a、MPEG2画像復号回路32a、MPEG2システムデータ復号回路36a、MPEG4サウンド復号回路28b、MPEG4画像復号回路32b、MPEG4システムデータ復号回路36bでそれぞれ復号される。MPEG2サウンド復号回路28a、MPEG2画像復号回路32aおよびMPEG2システムデータ復号回路36aをまとめてMPEG2復号回路(MPEG2デコーダ)を構成する。また、MPEG4サウンド復号回路28b、MPEG4画像復号回路32bおよびMPEG4システムデータ復号回路36bをまとめてMPEG4復号回路(MPEG4デコーダ)を構成する。なお、MPEG2の各データの復号方法および復号回路は、公知であるので、その説明を省略する。

【0113】MPEG4の具体的な復号方法および復号回路は先に説明した手順で行われるが、MPEG4画像オブジェクトの復号について補足する。MPEG4画像オブジェクトは、各画像オブジェクトにそれぞれ対応して復号を行う複数の同様の復号部を有するMPEG4画像復号回路32bにより復号される。ここでの復号方式は、既に説明したMPEG4の画像復号方式に基づくオブジェクト単位の復号であり、復号された画像データは、オブジェクトの数に相当する画像(v(1)からv(i))になる。

【0114】復号されたサウンドデータは、MPEG2サウンド復号回路28aおよびMPEG4サウンド復号回路28bの出力を多重または切換えるためのサウンド多重/切換回路52へ入力される。サウンド多重/切換回路52においては、MPEG2サウンドとMPEG4サウンドとの多重、または、どちらか一方を出力する切換えが行われるとともに、出力するサウンドの各種出力調整などが施される。サウンド多重/切換回路33は、サウンド制御部30に制御される。

【0115】サウンド制御部34は、シーン記述データ変換回路39からのMPEG4のシーン記述情報に応じて出力制御を行い、かつ、システムコントローラ38からのコマンドに応じて動作する。

【0116】システムコントローラ47は、指示入力部(操作部)45を介して入力されるユーザの指示により、複数のサウンドがある場合、どのサウンドを選択する

か、または、それらを多重するかなどの指示を受け付ける。そして、その指示およびMPEG4データ検出回路51の検出信号に応じたコマンド、つまりサウンドの多重/切換えを行わせるコマンドをサウンド制御部34に出力する。勿論、MPEG4サウンドオブジェクトがない場合はMPEG2サウンドオブジェクトのみの再生になる。サウンド多重/切換え回路52から出力される最終的なサウンドデータは、D/Aコンバータ29でアナログ信号に変換され、スピーカ31によりステレオ再生される。

【0117】次に、画像の再生について説明する。MPEG2画像復号回路32aおよびMPEG4画像復号回路32bの出力はシーン合成回路53に入力される。一方、システムコントローラ38は、MPEG4データ検出回路51の検出信号、および、指示入力部（操作部）45を介して入力されるユーザの指示に従いシーン合成のためのコマンドを表示制御部34へ出力する。なお、指示入力部45を介して入力されるユーザの指示には、合成表示させたいMPEG4画像オブジェクトを選択する指示などが含まれる。表示制御部34は、シーン記述データ変換回路39から入力されるMPEG4のシーン記述情報、および、システムコントローラ38から入力されるコマンドに応じて表示制御、つまりシーン合成回路53を制御する。

【0118】シーン合成回路53は、表示制御部34の制御に従い、MPEG2画像およびMPEG4画像のシーン合成を行う。なお、MPEG4画像オブジェクトのすべてを表示せずに、必要な画像オブジェクトだけを選択し合成し再生することが可能である。勿論、MPEG4画像オブジェクトがない場合はMPEG2画像のみの再生になる。合成された表示画像は、D/Aコンバータ33でアナログ信号に変換され、CRT35により表示される。もしくは、合成された表示画像を、ディジタル信号のまま液晶フラットディスプレイ(LCD)44などへ送り、表示させることもできる。

【0119】次に、システムデータの処理について説明する。MPEG2システムデータは、MPEG2システムデータ復号回路36aで復号され、MPEG2画像およびMPEG2サウンドを制御する各種コマンドとしてシステムコントローラ38へ入力される。システムコントローラ38は、必要に応じて、MPEG2システムデータを制御データとして利用する。

【0120】一方、MPEG4システムデータ（シーン記述情報を含む）は、MPEG4システムデータ復号回路36bで復号された後、復号されたシステムデータ中に含まれるシーン記述に関する情報がシーン記述データ変換回路39に入力される。その他のシステムデータは、MPEG4画像、MPEG4サウンド、シーン記述情報などを制御する各種コマンドとしてシステムコントローラ38へ入力され、必要に応じて制御データとして利用される。シーン記述データ変換回路39は、シーン記述情報に従い、MPEG4画像およびMPEG4サウンドの出力形態を規定するためのシーン記述データを、サウンド制御部30および表示制御部34へ出

力する。

【0121】ところで、ユーザ操作部45からは、前述したサウンドおよび画像の選択に関するユーザの指示のほかにも各種の指示が入力される。また、システムコントローラ38は、指示入力部45からの指示入力に従い、または、動作に応じた自動制御で本受信表示装置の各部の制御を統括していることは言うまでもない。

【0122】[データストリーム]ここで、テレビ情報としてのMPEG2データストリームにMPEG4データストリームを多重する方法を図20および図21を用いて説明する。

【0123】図20はMPEG4のデータストリーム構成を示す図である。図20に示すように、MPEG4のデータストリームには、オブジェクト1から5までのデータベースに、写真（自然）画像、サウンドを含むサウンドおよびコンピュータグラフィックス(CG)などの合成画像のオブジェクトなどが入る。さらに、MPEG4システムデータとして、表示出力制御のシーン記述情報(BIFS)と、その他必要な各種データ（サブデータ）が組み込まれる。

【0124】図21はMPEG2のデータストリームの伝送形式を示す、MPEG2トランスポートストリーム構造を示す図である。図21を用いて、MPEG4データストリームをMPEG2データストリームに多重する方法を説明する。

【0125】MPEG2トランスポートストリームは、固定長のトランスポートパケットによって多重され分離される。トランスポートパケットのデータ構造は、図21に示すように階層的に表され、それぞれ図21に示す項目を含む。それら項目を順に説明すると、8ビットの「同期信号(sync)」、パケット内のビットエラーの有無を示す「誤り表示（エラーインジケータ）」、このパケットのペイロードから新たなユニットが始まることを示す「ユニット開始表示」、このパケットの重要度を示す「プライオリティ（パケット優先度）」、個別ストリームの属性を示す「識別情報PID(Packet Identification)」、スクランブルの有無および種別を示す「スクランブル制御」、このパケットのアダプテーションフィールドの有無およびペイロードの有無を示す「アダプテーションフィールド制御」、同じPIDをもつパケットが途中で一部棄却されたかどうかを検出するための情報である「巡回カウンタ」、付加情報や、スタッフィングバイトをアダプテーションフィールドに入れることができる「アダプテーションフィールド」、並びに、ペイロード（画像やサウンドの情報）である。アダプテーションフィールドは、フィールド長、その他の個別ストリームに関する各種項目、オプショナルフィールド、並びに、スタッフィングバイト（無効データバイト）からなる。

【0126】本実施形態においては、テレビ情報のサブ画像またはサウンドデータとしてのMPEG4のデータストリーム、および、それを識別するためのIDをオプショナルフィールドにおける付加データの一つと見做し、オプショナルフィールド内に多重する。

【0127】つまり、メインであるテレビ情報の構成はMPEG2データストリーム（トランスポートストリーム）である。そして、図21に一例を示すように、データ量としては微かな写真（自然）画像、CG、文字などの画像オブジェクト（オブジェクトAおよびB）、サウンドオブジェクト（オブジェクトC）、シーン記述情報（BIFS）、並びに、その他必要なデータ（サブデータ）を組み合わせたMPEG4データストリームを構成する。このMPEG4データストリームを、MPEG2のシステムデータ中のオプションフィールドの一部として多重させることで、MPEG2とMPEG4とのデータストリーム多重伝送が実現される。

【0128】〔再生表示例〕次に、本実施形態の受信表示装置による再生表示例を図22から図26を参照して説明する。なお、図22から図26に示すのはMPEG2により放送される野球中継番組を例とするものである。

【0129】図22は基本映像100としてのMPEG2放送画像のみが表示されたシーンの例である。図23は基本映像100にMPEG4画像オブジェクトである試合経過情報101がシーン合成された例である。図24は、試合経過情報101に加えてMPEG4画像オブジェクトであるプレイバック映像102がシーン合成され、MPEG4サウンドオブジェクトであるプレイバック映像のサウンドが多重された例である。図25は基本映像103にMPEG4画像オブジェクトである選手情報104がシーン合成された例である。図26は基本映像100にMPEG4画像オブジェクトである天気予報105およびニュース106がシーン合成された例である。

【0130】このように、本実施形態によれば、MPEG2のどのような映像（画像）に対してもMPEG4オブジェクトを多重（シーン合成を含む）することができ、多重された映像（画像）またはサウンドを表示または再生（出力）することができる。

【0131】また、多重されるMPEG4オブジェクトが画像オブジェクトの場合も、静止画像に限らず、リアルタイム動画やそれに附随するサウンドなども扱うことができる。また、眼の不自由な人への副サウンドにも利用できる。

【0132】さらに、図示したように、MPEG2画像オブジェクトであるメイン画像（映像）の内容に関する詳細情報をMPEG4画像オブジェクトとして提供したり、ユーザが望むであろう過去のシーンを随時提供することなども可能である。また、メイン画像（映像）とは関係のない天気予報、道路情報およびニュースなどの生活情報などもMPEG4画像オブジェクトとして提供することができ、様々な用途への適用が期待できる。

【0133】〔動作手順〕図27は本実施形態のデジタルテレビ受信表示装置の動作手順を説明するためのフローチャートである。

【0134】MPEG2のデジタルテレビ情報が放送衛星もしくはケーブルから受信され（ステップS1）、チューナ23または24を用いて、受信されたデジタルテレビ情

報から番組が選局される（ステップS2）。

【0135】続いて、MPEG4データ検出回路51により、選局された番組のMPEG2データストリームに多重されているサブテレビ情報のMPEG4データの検出が行われ（ステップS3）、その検出結果に基づき、MPEG2データストリームにMPEG4データが含まれているかが判断される（ステップS4）。MPEG4データが含まれていない場合は、受信されたMPEG2データのみが、前述した処理によりサウンド、画像およびシステムデータに分けられて復号される（ステップS5）。

【0136】一方、MPEG4データが含まれている場合は、MPEG2データストリームからMPEG4データストリームが分離され、MPEG2データおよびMPEG4データそれぞれが、前述した処理によりサウンド、画像およびシステムデータに分けられて復号される（ステップS6）。さらに、MPEG2の映像（画像）およびサウンドと、MPEG4のシーンおよびサウンドとの出力形態が、シーン合成回路53およびサウンド多重/切換回路52などにより設定される（ステップS7）。

【0137】こうして、ステップS5で復号されたMPEG2の映像（画像）およびサウンド、あるいは、ステップS7でMPEG2の映像（画像）にMPEG4の画像（映像）が合成されたシーンおよびMPEG2サウンドとMPEG4サウンドとが多重/切換えされたサウンドが表示および再生される（ステップS8）。

【0138】なお、必要に応じて、ステップS1からS8の処理またはその一部が繰り返されることは言うまでもない。

【0139】このように、本実施形態によれば、メインのテレビ情報であるMPEG2のデータストリームへ、サブ情報である画像（映像）やサウンドのMPEG4データが多重されたデータストリームからなるデジタルテレビ放送を受信し、映像（画像）およびサウンドを再生することができる。従って、より多機能なデータ伝送であるデジタルテレビ放送を実現するとともに、視聴者であるユーザにより便利なテレビ番組の表示および再生が可能になる。

【0140】また、サブ情報の伝送にMPEG4を用いることにより、現行のデジタルテレビ放送システムであるMPEG2との親和性を高めることも容易であり、MPEG2用の既存コンテンツを有効に活用することができる。

【0141】また、画像（映像）や音声を含むサウンドをオブジェクト単位で扱うことができるMPEG4は、サブ情報を伝送する場合に最適なデータ形態であるといえる。

【0142】〔その他〕図28に示すMPEG2のみに対応したデジタルテレビ受信表示装置が、上述したメインのテレビ情報のMPEG2データストリームにサブテレビ情報のMPEG4データが多重されたデジタルテレビ放送を受信した場合について説明する。

【0143】図28に示すMPEG2デコーダ61は、MPEG2で符号化された画像、サウンドおよびシステムデータを復号する。また、サウンド制御部62は、復号されたMPEG2サウンドの再生を制御し、表示制御部63は、復号されたMPEG2画像（映像）の再生表示を制御する。図28に示すテレビ放送受信表示装置はMPEG4のデコード機能を備えていないので、MPEG2データストリームにサブ情報として多重されたMPEG4データを復号することはできず、再生できるのはMPEG2の画像（映像）およびサウンドだけである。

【0144】この場合、MPEG2データストリームは図21に示したように構成され、MPEG2データストリーム中のオプションフィールドにMPEG4データストリームが含まれている。図28に示すテレビ放送受信表示装置のMPEG2デコーダ61は、その復号処理においてMPEG4データを無視するようにする。

【0145】このようにすれば、図21に示すMPEG2データストリーム、つまりMPEG4データストリームが多重されたデジタルテレビ放送を受信する場合、MPEG4の復号および再生機能を有する受信表示装置は、MPEG2およびMPEG4両方のテレビ放送情報を復号および再生することができる。また、MPEG4の復号および再生機能を備えない受信表示装置は、基本のMPEG2データストリームであるMPEG2のテレビ放送情報だけを復号および再生することができる。

【0146】このように、MPEG2だけに対応する受信表示装置に対しても、上述したメインのテレビ情報のMPEG2データストリームにサブテレビ情報のMPEG4データが多重されたデジタルテレビ放送のデータ伝送方式は対応することができる。従って、上述したデジタルテレビ放送のデータ伝送方式は、受信表示装置（テレビ受像機）の機能に関わらず放送することができ、受信表示装置がMPEG4対応へ移行する期間においても導入が可能である。

【0147】

【第2実施形態】以下、本発明にかかる第2実施形態の映像サウンド再生表示装置を説明する。なお、本実施形態において、第1実施形態と略同様の構成については、同一符号を付して、その詳細説明を省略する。

【0148】第2実施形態では、MPEG2で符号化されたデータを扱うDVDなどのパッケージメディアに、上述したMPEG2データとMPEG4データとの多重が利用できることを説明する。

【0149】図29はMPEG2で符号化されたデータを扱うDVDなどのパッケージメディア用の再生表示装置である。

【0150】蓄積メディア81は、デジタルビデオデータが保持された記録媒体である。蓄積メディア81に記録されたデジタルビデオデータは、MPEG2で符号化されたメインビデオ情報に、MPEG4で符号化されたサブビデオ情報が多重されたデータ形態で記録されている。この

MPEG2とMPEG4との多重には、図21により説明した多重方法が採用される。

【0151】図29において、蓄積メディア81に記録されたデジタルビデオデータは、再生処理回路82により再生され、誤り訂正回路83により誤り訂正が行われた後、MPEG4データ検出回路29へ送られる。以下、図19を用いて説明したのと同様の手順により、画像、サウンドおよびシステムデータが復号され、映像（画像）およびサウンドが表示および再生される。

【0152】図30は本実施形態の再生表示装置の動作手順を説明するためのフローチャートである。

【0153】MPEG2のデジタルテレビ情報が蓄積メディア81から再生され（ステップS11）、MPEG4データ検出回路51により、再生されたビデオ情報のMPEG2データストリームに多重されているサブビデオ情報のMPEG4データの検出が行われ（ステップS12）、その検出結果に基づき、MPEG2データストリームにMPEG4データが含まれているかが判断される（ステップS13）。MPEG4データが含まれていない場合は、再生されたMPEG2データのみが、前述した処理によりサウンド、画像およびシステムデータに分けられて復号される（ステップS14）。

【0154】一方、MPEG4データが含まれている場合は、MPEG2データストリームからMPEG4データストリームが分離され、MPEG2データおよびMPEG4データそれぞれが、前述した処理によりサウンド、画像およびシステムデータに分けられて復号される（ステップS15）。さらに、MPEG2の映像（画像）およびサウンドと、MPEG4のシーンおよびサウンドとの出力形態が、シーン合成回路53およびサウンド多重/切換回路52などにより設定される（ステップS16）。

【0155】こうして、ステップS14で復号されたMPEG2の映像（画像）およびサウンド、あるいは、ステップS16でMPEG2の映像（画像）にMPEG4の画像（映像）が合成されたシーンおよびMPEG2サウンドとMPEG4サウンドとが多重/切換えされたサウンドが表示および再生される（ステップS17）。

【0156】なお、必要に応じて、ステップS11からS17の処理またはその一部が繰り返されることは言うまでもない。

【0157】なお、図29に示す再生表示装置は、図19および図29に示すデジタルインタフェイス(DIF)54を介して、図19に示すデジタルテレビ放送受信表示装置へビデオデータを伝送することができる。

【0158】このように、上述したMPEG2とMPEG4とを多重したデータストリームによりメイン情報およびサブ情報の画像（映像）、サウンドおよびシステムデータを伝送する技術は、第1実施形態のデジタルテレビ放送受信表示装置に限らず、DVDのような蓄積メディア、および、その蓄積メディアを活用する再生表示装置にも適用することができる。

【0159】

【他の実施形態】なお、本発明は、複数の機器（例えばホストコンピュータ、インタフェイス機器、リーダー、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【0160】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0161】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0162】

【発明の効果】以上説明したように、本発明によれば、メイン画像に関係する、または関係しなくとも、その時々にも所望される情報を少なくとも画像の形態で再生する機能を提供することができる。

【図面の簡単な説明】

【図1】衛星放送を用いたデジタル放送受信機の構成例を示すブロック図、

【図2】複数種類のオブジェクトを同時に入力し符号化処理する構成例を示すブロック図、

【図3】ユーザ操作（編集）を考慮に入れたシステムの構成例を示す図、

【図4】ビデオオブジェクトに関するVOP処理回路のエ

ンコーダ側のブロック図、

【図5】ビデオオブジェクトに関するVOP処理回路のデコーダ側のブロック図、

【図6】VOPの符号化および復号の全体構成を示すブロック図、

【図7A】VOPを構成する情報を示す図、

【図7B】VOPを構成する情報を示す図、

【図8】テキストチャ符号化のAC/DC予測符号化を説明するための図、

【図9A】スケーラビリティを実現するためのシンタックスの階層構造を説明するための図、

【図9B】スケーラビリティを実現するためのシンタックスの階層構造を説明するための図、

【図10A】ワーブを説明する図、

【図10B】ワーブの種類を説明する図、

【図11】ワーブを説明する図、

【図12】シーン記述情報の構成例を示す図、

【図13】MPEG4オーディオの符号化方式の種類を示す図、

【図14】オーディオ符号化方式の構成を示す図、

【図15】MPEG4のシステム構造を説明する図、

【図16】MPEG4のレイヤ構造を説明する図、

【図17】双方向復号を説明する図、

【図18】重要情報の複数回伝送を説明する図、

【図19】本発明にかかる第1実施形態のテレビ放送受信装置の構成例を示すブロック図、

【図20】MPEG2データストリームにMPEG4データストリームを多重する方法を説明する図、

【図21】MPEG2データストリームにMPEG4データストリームを多重する方法を説明する図、

【図22】再生表示例を説明する図、

【図23】再生表示例を説明する図、

【図24】再生表示例を説明する図、

【図25】再生表示例を説明する図、

【図26】再生表示例を説明する図、

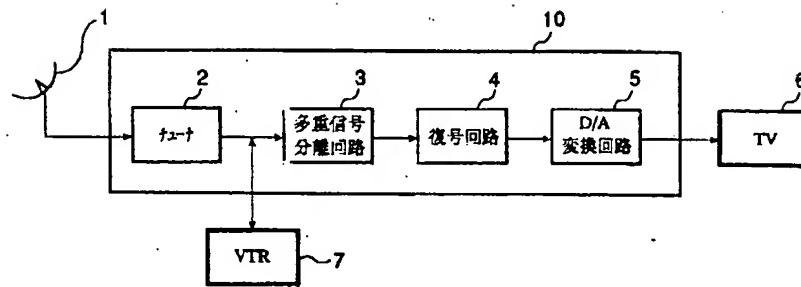
【図27】デジタルテレビ受信表示装置の動作手順を説明するためのフローチャート、

【図28】MPEG2のみに対応したデジタルテレビ受信表示装置の構成例を示す図、

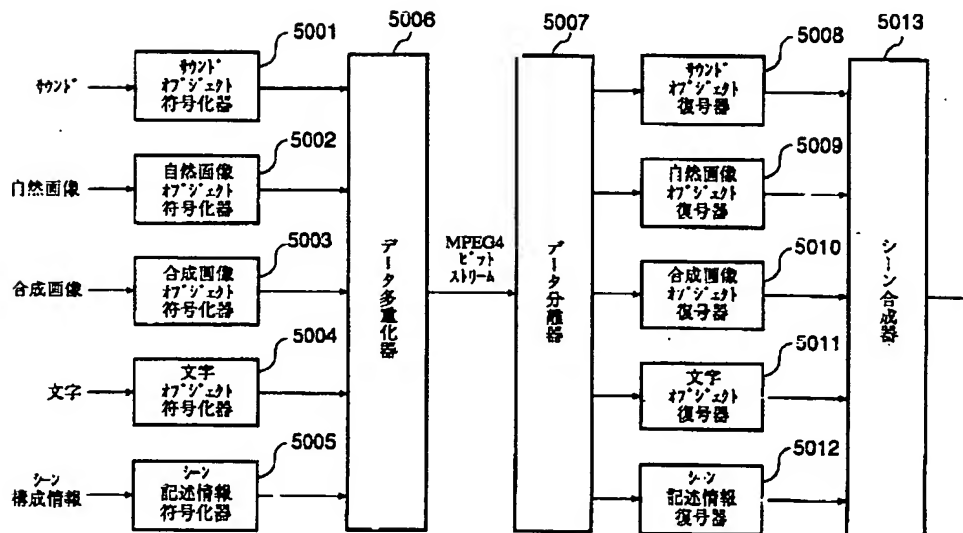
【図29】第2実施形態のパッケージメディア用の再生表示装置の構成例を示す図、

【図30】再生表示装置の動作手順を説明するためのフローチャートである。

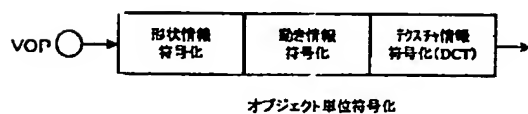
【図1】



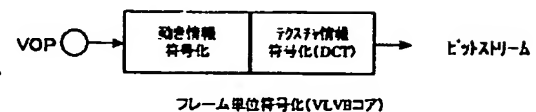
【図2】



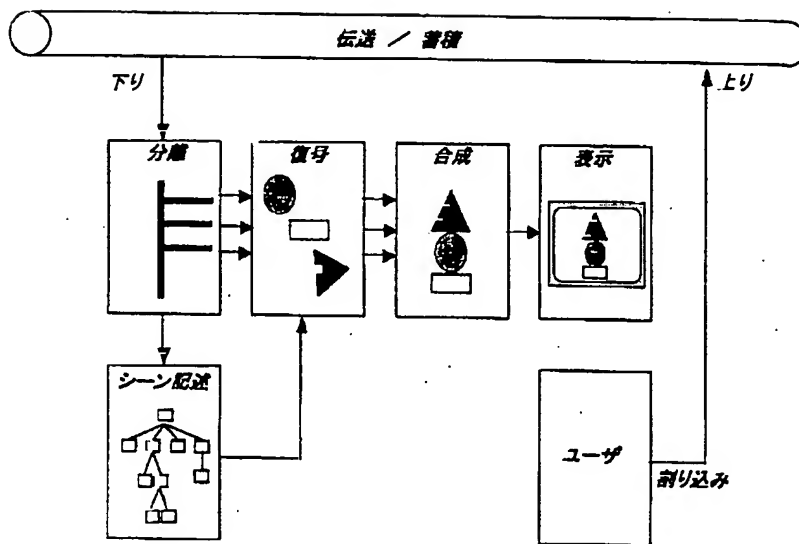
【図7A】



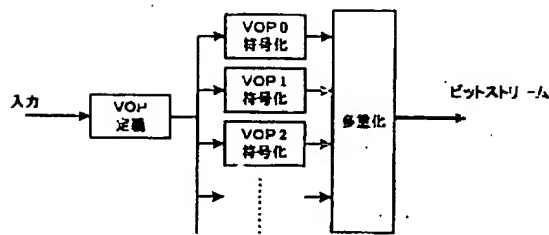
【図7B】



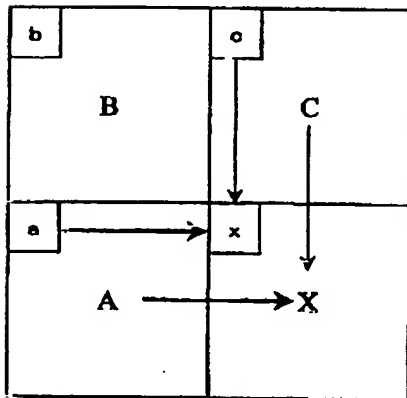
【図3】



【図4】

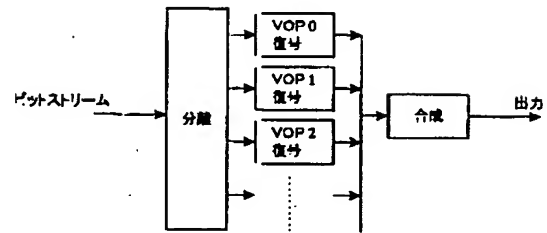


【図8】

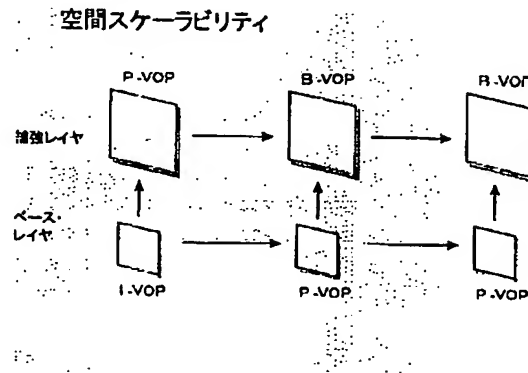


a,b,c,x: 直流成分の量子化係数
A,B,C,X: 交流成分の量子化係数

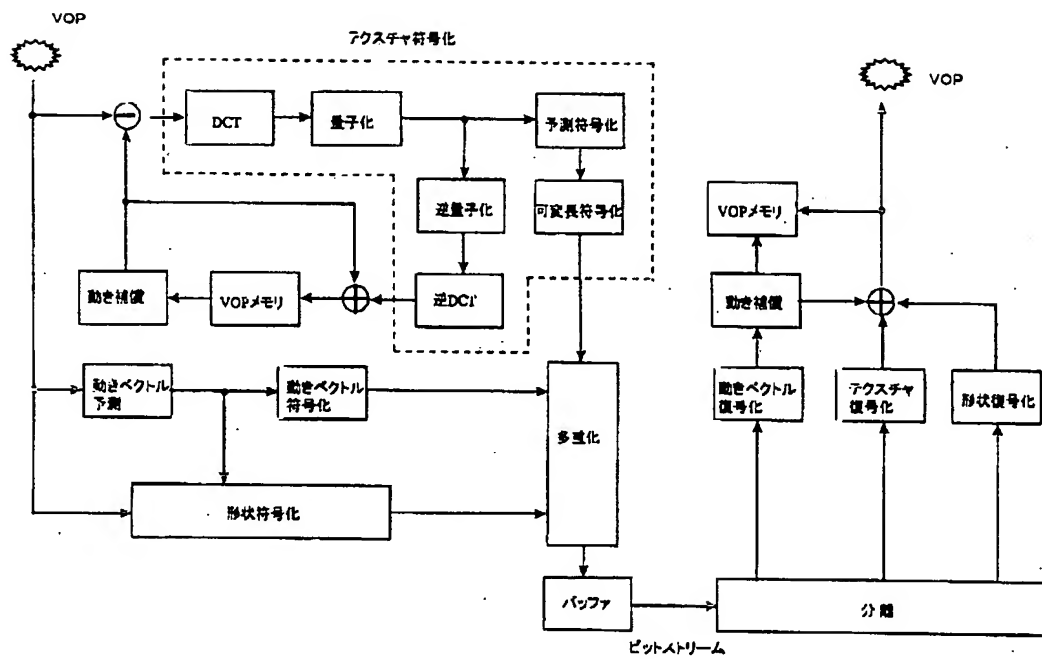
【図5】



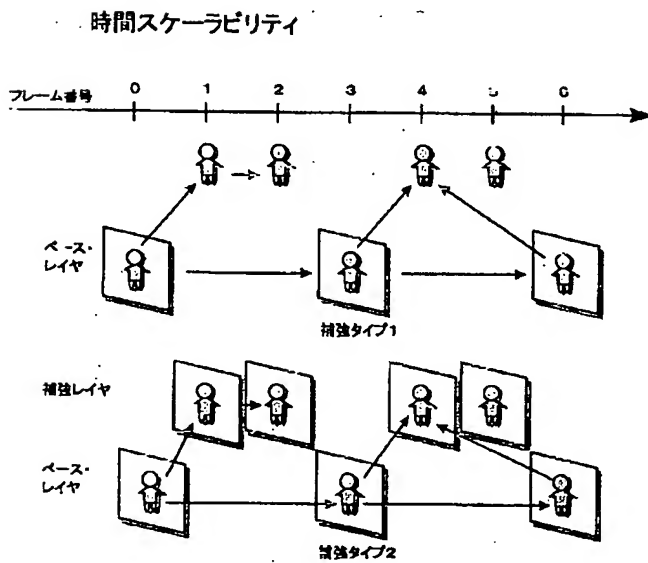
【図9B】



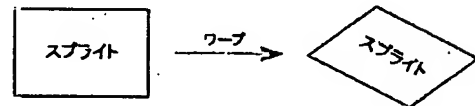
【図6】



【図9A】



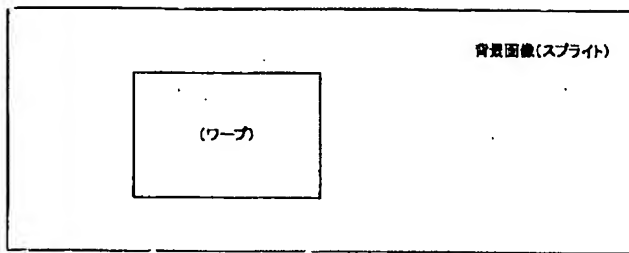
【図10A】



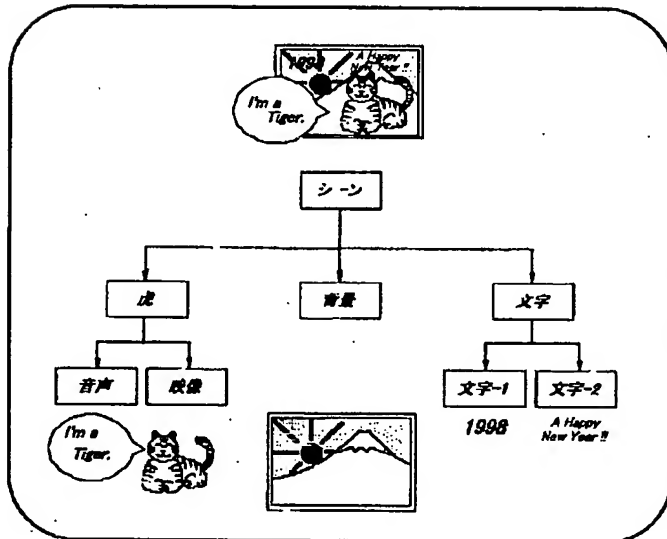
【図10B】

遠近法変換	$x' = (ax + by + c) / (gx + hy + 1)$ $y' = (dx + ey + f) / (gx + hy + 1)$
アフィン変換	$x' = ax + by + c$ $y' = dx + ey + f$
等方位大(a)/回転(θ)/移動(c,f)	$x' = \cos \theta x + \sin \theta y + c$ $y' = -\sin \theta x + \cos \theta y + f$
平行移動	$x' = x + c$ $y' = y + f$

【図11】



【図12】

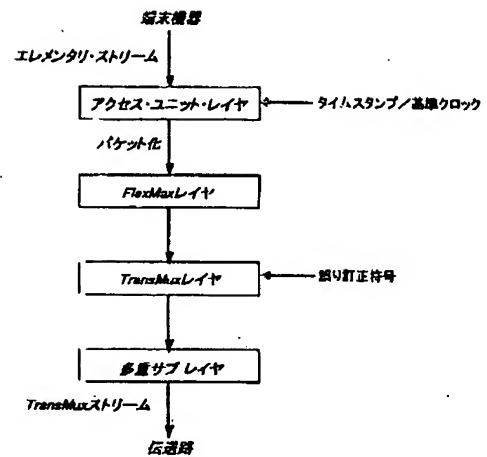


【図13】

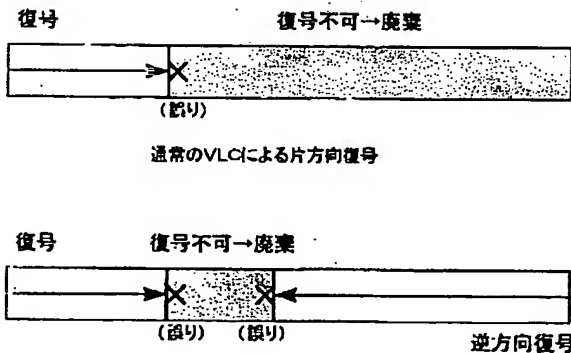
符号化方式		ビットレート kbit/s
ハフマン符号化	IL	8-18
	HVXC	2-6
CELP符号化	WB-CELP	14-24
	NB-CELP	4-12
時間/周波数変換符号化 (T/F変換)	AAC準拠	24-64
	TwinVQ	8-40
SNHC	SA符号化(音声合成)	-
	TTS符号化(音声合成)	-

CELP: Code Excited Linear Prediction
SNHC: Synthetic Natural Hybrid Coding

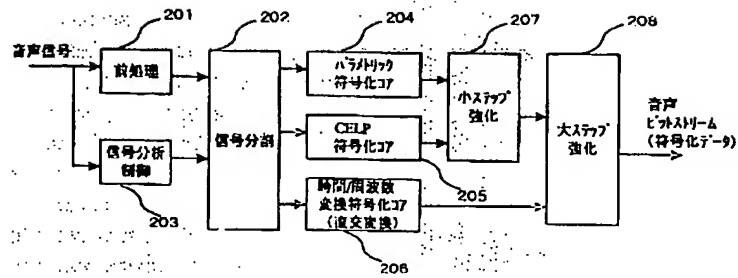
【図15】



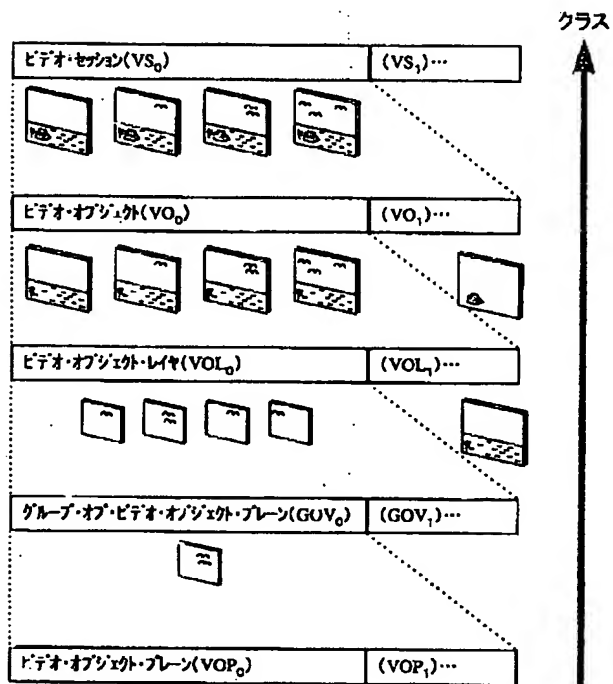
【図17】



【図14】



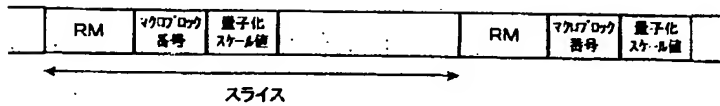
【図16】



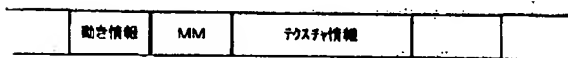
VS: Video Session
 VO: Video Object
 VOL: Video Object Layer
 GOV: Group Of Video Object Plane
 VOP: Video Object Plane

【図18】

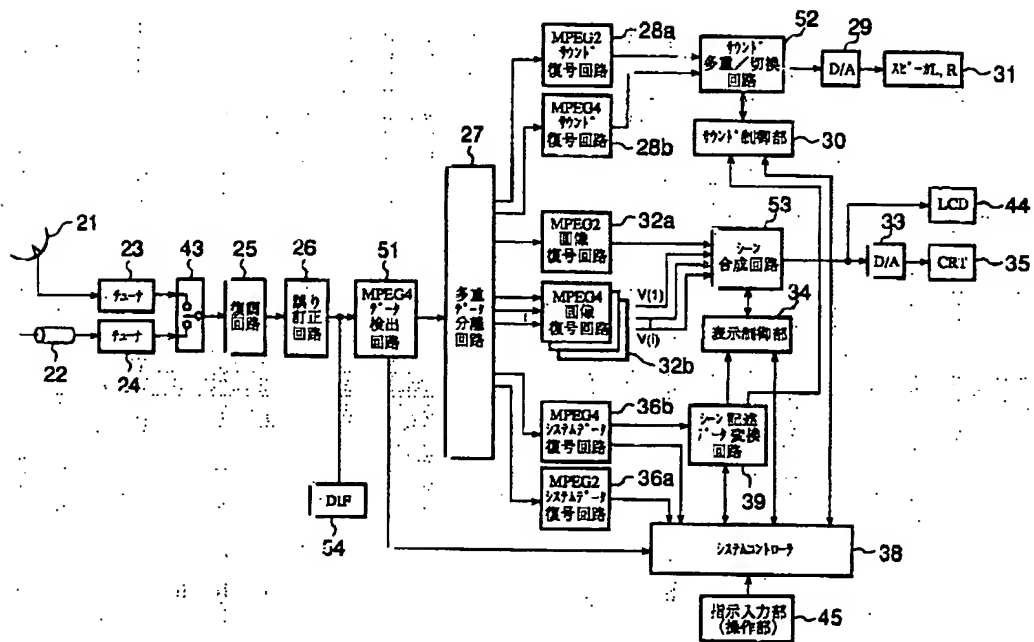
RM(再同期マーカ)の挿入



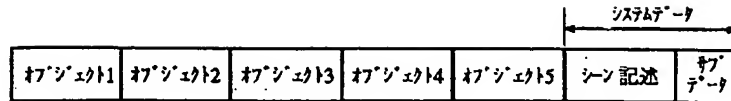
MM(動きマーカ)の挿入<データ・パーティショニング>



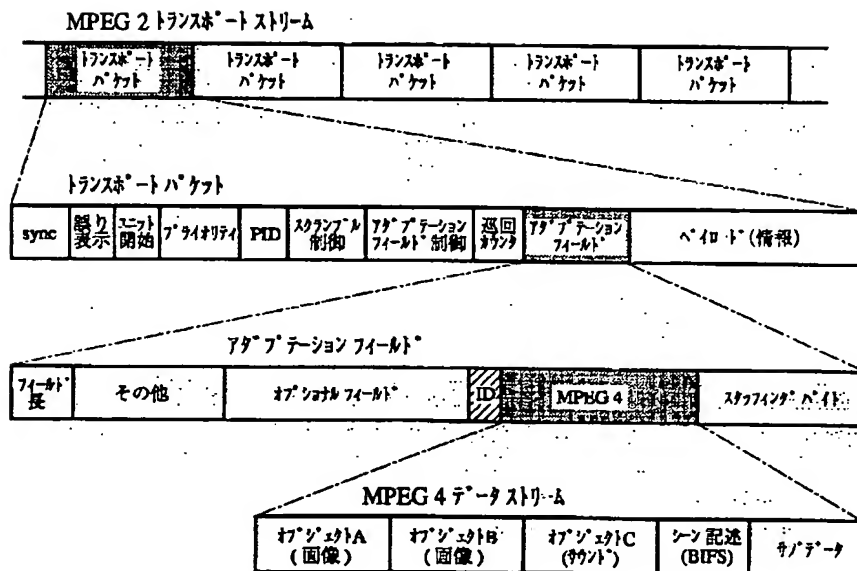
【図19】



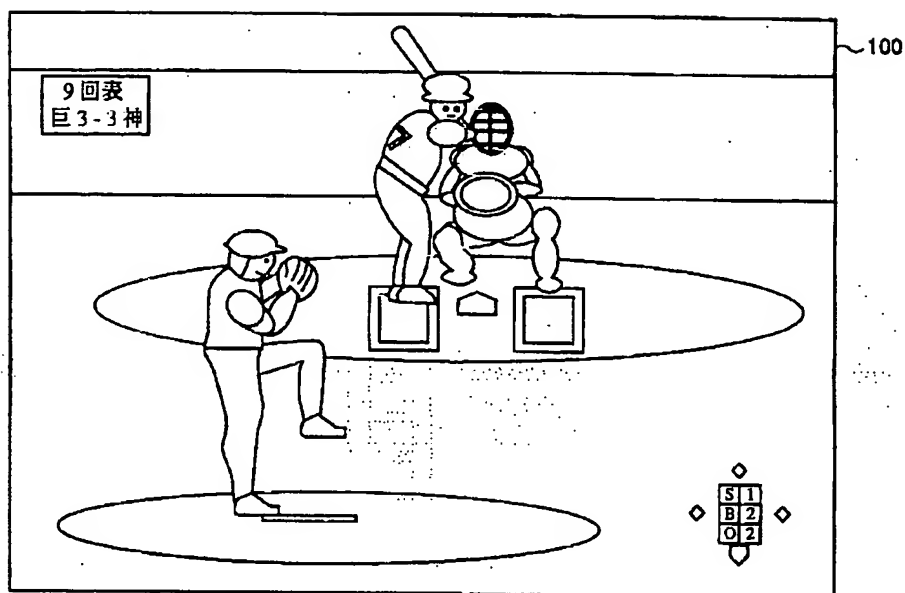
【図20】



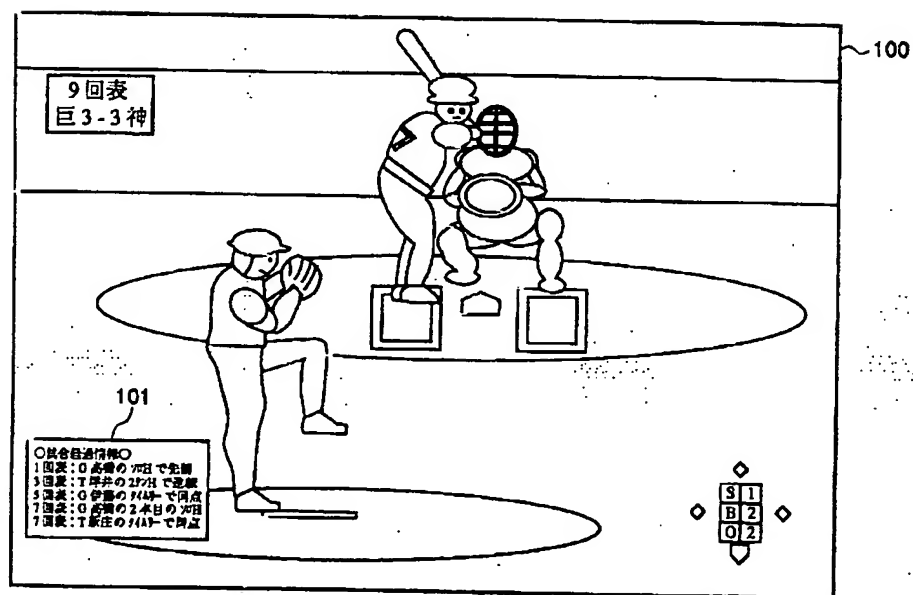
【図21】



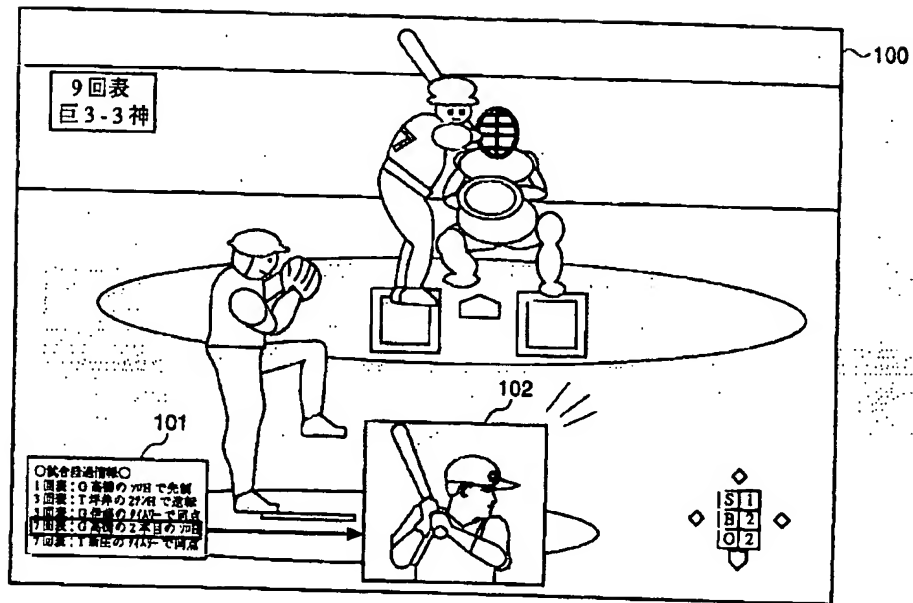
【図22】



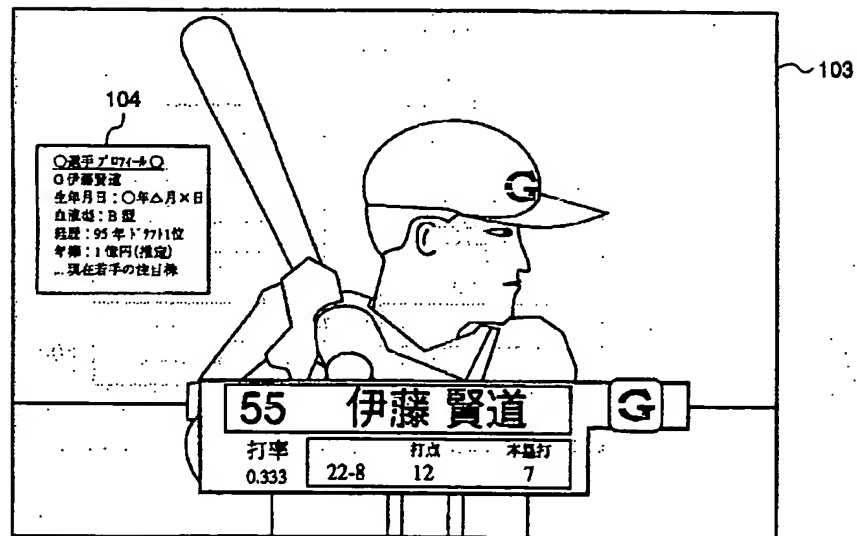
【図23】



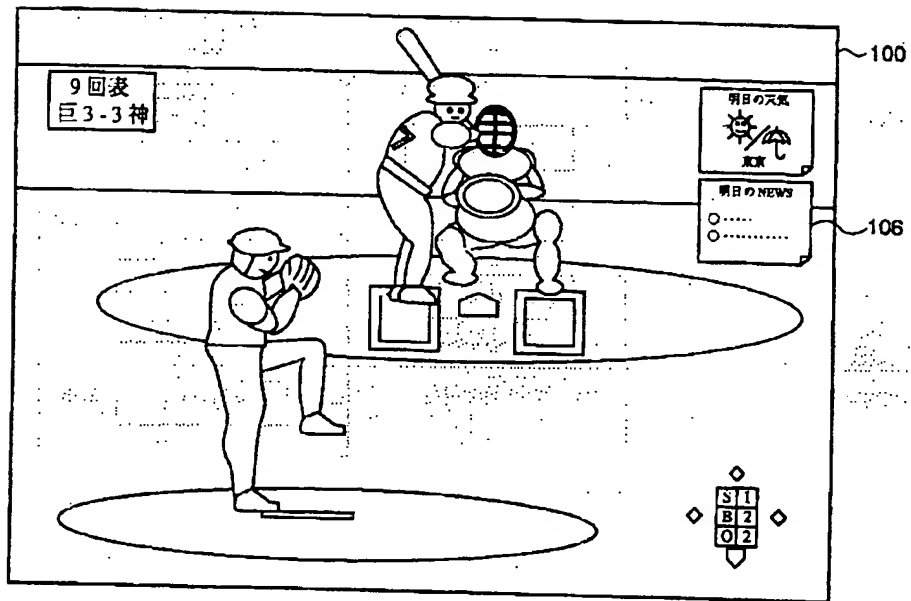
【図24】



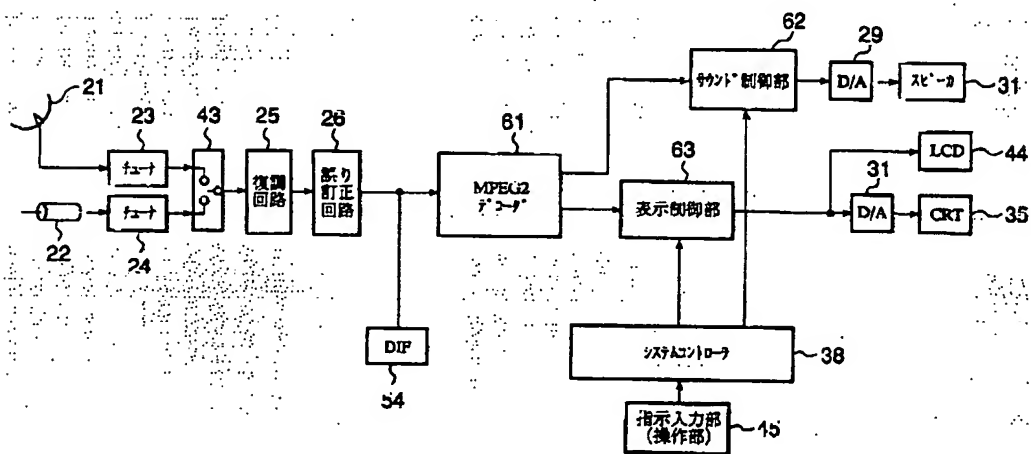
【図25】



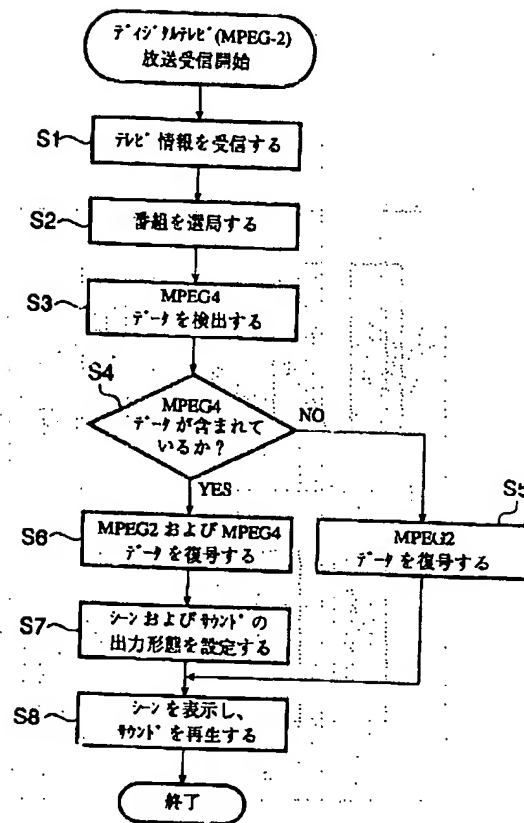
【図26】

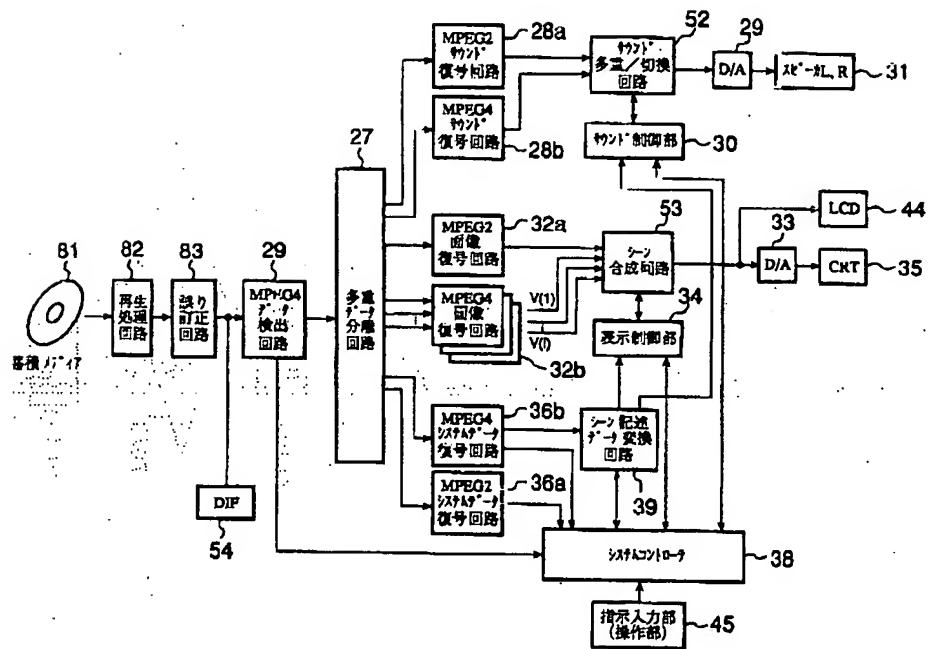


【図28】

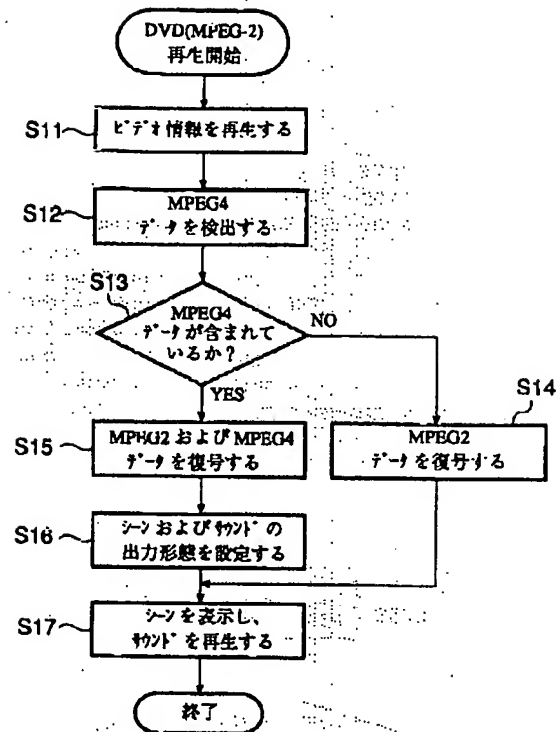


【図27】





【図30】



フロントページの続き

Fターム(参考) 5C059 MA00 MA05 MA10 MA14 MA23
 MA24 MA32 MA33 MB00 MB04
 MB14 MB22 MB23 MB24 MB27
 MC01 MC06 MC11 MC33 MC35
 MC38 ME01 ME10 NN11 PP19
 RB02 RB03 RB10 RB16 RC04
 RC07 RC16 RC24 RC31 RC32
 RC37 SS02 SS12 SS13 SS26
 UA05 UA22 UA39